

University of Groningen

## Tagging *Hansenula polymorpha* genes by random integration of linear DNA fragments (RALF)

van Dijk, R; Faber, KN; Hammond, AT; Glick, BS; Veenhuis, M; Kiel, JAKW

*Published in:*  
MOLECULAR GENETICS AND GENOMICS

*DOI:*  
[10.1007/s004380100584](https://doi.org/10.1007/s004380100584)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2001

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

van Dijk, R., Faber, KN., Hammond, AT., Glick, BS., Veenhuis, M., & Kiel, JAKW. (2001). Tagging *Hansenula polymorpha* genes by random integration of linear DNA fragments (RALF). *MOLECULAR GENETICS AND GENOMICS*, 266(4), 646-656. <https://doi.org/10.1007/s004380100584>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

R. van Dijk · K. N. Faber · A. T. Hammond  
B. S. Glick · M. Veenhuis · J. A. K. W. Kiel

## Tagging *Hansenula polymorpha* genes by random integration of linear DNA fragments (RALF)

Received: 4 April 2001 / Accepted: 27 August 2001 / Published online: 6 October 2001  
© Springer-Verlag 2001

**Abstract** We have investigated the feasibility of using gene tagging by restriction enzyme-mediated integration (REMI) to isolate mutants in *Hansenula polymorpha*. A plasmid that cannot replicate in *H. polymorpha* and contains a dominant zeocin resistance cassette, pREMI-Z, was used as the integrative/mutagenic plasmid. We observed that high transformation efficiency was primarily dependent on the use of linearised pREMI-Z, and that the addition of restriction endonuclease to linearised pREMI-Z prior to transformation increased the transformation frequency only slightly. Integration of linearised pREMI-Z occurred at random in the *H. polymorpha* genome. Therefore, we termed this method Random integration of Linear DNA Fragments (RALF). To explore the potential of RALF in *H. polymorpha*, we screened a collection of pREMI-Z transformants for mutants affected in peroxisome biogenesis (*pex*) or selective peroxisome degradation (*pdd*). Many previously described *PEX* genes were obtained from the mutant collection, as well as a number of new genes, including *H. polymorpha* *PEX12* and genes whose function in peroxisome biogenesis is still unclear. These results demonstrate that RALF is a powerful tool for tagging genes in *H. polymorpha* that should make it possible to carry out genome-wide mutagenesis screens.

**Keywords** Gene tagging · Genomic integration · Mutants · Peroxisome homeostasis · Yeast

Communicated by C. P. Hollenberg

R. van Dijk · K.N. Faber · M. Veenhuis · J.A.K.W. Kiel (✉)  
Eukaryotic Microbiology, Groningen Biomolecular Sciences  
and Biotechnology Institute (GBB), University of Groningen,  
Kerklaan 30, 9751 NN Haren, The Netherlands  
E-mail: J.A.K.W.Kiel@biol.rug.nl  
Tel.: +31-50-3632218  
Fax: +31-50-3632154

A.T. Hammond · B.S. Glick  
Department of Molecular Genetics and Cell Biology,  
University of Chicago, 920 East 58<sup>th</sup> Street,  
Chicago, IL 60637, USA

### Introduction

In recent years, the methylotrophic yeast *Hansenula polymorpha* has become increasingly important for both commercial and scientific applications (van Dijk et al. 2000). The availability of advanced genetic tools is an essential prerequisite for the full exploitation of the potential of this organism. To date, novel genes involved in a specific cellular process can only be identified through the complementation of specific mutants obtained by N-methyl-N'-nitro-nitrosoguanidine (NTG) treatment or UV irradiation. Although mutants are easy to generate in *H. polymorpha*, functional complementation and the identification of the responsible genetic elements is time consuming and laborious. On the other hand, targeted integration of plasmid DNA into the *H. polymorpha* genome is a well established technique (Faber et al. 1992). Interestingly, in addition to site-specific integration, non-specific integration of plasmid DNA was observed quite frequently in such studies. The frequency of non-specific integration was highly dependent on the plasmid used. When little or no homology existed between the plasmid and the genome of *H. polymorpha*, a higher tendency towards illegitimate integration was observed (Faber et al. 1992). Therefore, insertional mutagenesis may be a useful strategy for generating mutants in *H. polymorpha*.

In bacteria, insertional mutagenesis strategies generally involve the use of mobile genetic elements (Manoil and Traxler 2000). Unfortunately, these elements are not available for *H. polymorpha*. However, an insertional mutagenesis procedure for *Saccharomyces cerevisiae* has been described that utilises the in vivo action of restriction endonucleases (Schiestl and Petes 1991). This method was termed Restriction Enzyme-Mediated Integration (REMI) (Kuspa and Loomis 1992), and involves the addition of a restriction enzyme to linearised plasmid DNA in transformation mixtures. It is assumed that the in vivo action of the added enzyme generates free chromosomal ends in the genome of the transformants, which are ligated to the linearised plasmid ends by cellular DNA repair

mechanisms, ultimately resulting in integration of the plasmid DNA into the host genome, a process which is expected to generate mutants. The ability to recover the integrated plasmids, together with their flanking genomic regions, from the mutants by plasmid rescue in *Escherichia coli* greatly enhances the usefulness of this insertional mutagenesis procedure. To date, REMI has been successfully used to generate mutant strains in several organisms, including *Dictyostelium* sp. (Kuspa and Loomis 1992), *Aspergillus* sp. (Shuster and Bindel 1999), *Ustilago maydis* (Bolker et al. 1995), *Coprinus cinereus* (Granado et al. 1997), *Lentinus edodes* (Sato et al. 1998), *Polysphondylium pallidum* (Kawabe et al. 1999), *Gibberella fujikuroi* (Linnemann et al. 1999) and *Candida albicans* (Brown et al. 1996).

Here, we have investigated the use of REMI in the methylotrophic yeast *H. polymorpha*. Upon linearisation, the plasmid pREMI-Z integrated at high frequency into the genome of *H. polymorpha*. Our data show that these integration events occurred randomly and did not require the addition of restriction enzyme prior to transformation. Therefore, we designated this process Random Integration of Linear DNA Fragments (RALF). The use of RALF to mutagenise and tag genes in *H. polymorpha* was evaluated by screening for mutants that were affected in either the biogenesis or the selective degradation of peroxisomes, organelles that are indispensable for growth of *H. polymorpha* on methanol as sole carbon and energy source (van der Klei et al. 1991). Sequence analysis of the genomic regions flanking the integrated pREMI-Z cassettes in selected mutants indicated that these mutants were affected in previously identified as well as novel genes involved in peroxisome biogenesis. The isolation of these mutants clearly demonstrates the application of RALF as a mutagenic procedure in *H. polymorpha*.

## Materials and Methods

### Organisms and growth conditions

The *E. coli* strains DH5 $\alpha$  and XL1-Blue were cultivated as described (Sambrook et al. 1989), except that zeocin (25  $\mu$ g/ml) (Invitrogen) selection was performed on LB medium containing 0.5% NaCl. The following *H. polymorpha* strains were used in this study: NCYC495 (*leu1.1*), NCYC495 (*leu1.1 ura3*) (Gleeson and Sudbery 1988) and HF246 (NCYC495::[P<sub>AOX</sub> eGFP-SKL] (*leu1.1*) (this study)). *H. polymorpha* cells were grown in batch cultures at 37°C on (1) selective minimal medium containing 0.67% (w/v) Yeast Nitrogen Base without amino acids (Difco), supplemented with 1% (w/v) glucose (YND) or 0.5% (v/v) methanol (YNM), or (2) on rich medium containing 1% (w/v) yeast extract, 1% (w/v) peptone and 1% (w/v) glucose (YPD), or (3) mineral medium (van Dijken et al. 1976) supplemented with 0.5% (w/v) glucose, 0.5% (v/v) methanol or a mixture of 0.1% (v/v) glycerol and 0.5% (v/v) methanol as carbon source, together with 0.25% (w/v) ammonium sulphate as nitrogen source. When required, amino acids and nucleotides (30  $\mu$ g/ml) and zeocin (100  $\mu$ g/ml) were added. For growth on plates, the media were supplemented with 1.5% granulated agar.

### Plasmid and reporter strain construction

Plasmid pREMI-Z was constructed as follows. Two 68mer oligonucleotide primers (5'-GATCGGAAACAGCTATGACCATG

TCAGTCAGTCAGGATCCTAGCTAGCTAGACTGGCCGTC-GTTTTAC-3' and 5'-GATCGTAAAACGACGGCCAGTCTAGCTAGTAGGATCCTGACTGACTGACATGGTCATAGCTGTTTCC-3') were synthesised, and annealed to yield an adapter containing the M13/pUC reverse and universal sequencing primers, and a central *Bam*HI restriction site, as well as stop codons in all six frames. Vector pPICZ- $\alpha$  (Invitrogen) was digested with *Bam*HI and *Bgl*II, and the resulting 1.9-kb fragment which contains the zeocin resistance cassette and the ColE1 origin of replication was ligated to the annealed adapter, resulting in pREMI-Z. The complete construct was verified by sequencing (see Genbank Accession No. AF282723). The *H. polymorpha* expression plasmid pFEM156, containing the eGFP.SKL reporter gene under the control of the *H. polymorpha* alcohol oxidase promoter (P<sub>AOX</sub>) was constructed by inserting a 2.6-kb *Eco*RV-*Bgl*II fragment from pFEM34 (R. van Dijk et al., submitted), containing the P<sub>AOX</sub>-eGFP-SKL expression cassette, into *Asp*718 (Klenow-treated) + *Bam*HI-digested pH11 (Kiel et al. 1999), a *URA3*-based *H. polymorpha* integration vector. For the construction of *H. polymorpha* HF246, the strain NCYC495 (*leu1.1 ura3*) was transformed with *Stu*I-linearised pFEM156. Site-specific integration of the linear plasmid into the genomic *AOX* locus and selection of a single-copy integrant was performed as described by Faber et al. (1994a).

### Molecular techniques

Genetic manipulations of *H. polymorpha* were performed as described previously (Gleeson and Sudbery 1988; Faber et al. 1992, 1994a). Standard recombinant DNA techniques were carried out essentially according to Sambrook et al. (1989). Transformation of *H. polymorpha* was performed by electroporation as described previously (Faber et al. 1994b). pREMI-Z transformations were performed similarly, except that increasing amounts of the restriction enzyme *Bam*HI were added to *Bam*HI-linearised pREMI-Z prior to electroporation. Control transformations were done with buffer of the same composition. Southern analysis was performed using the ECL Direct nucleic acid labeling and detection system (Amersham). Restriction enzymes and biochemicals were obtained from Boehringer Mannheim and used as detailed by the manufacturer. To clone the genomic fragment containing the full-length *PEX12* ORF, the Mut<sup>-</sup> mutant RVD3 was transformed with a *H. polymorpha* genomic library constructed in the vector pYT3 (Tan et al. 1995). Leucine prototrophic transformants were screened on YNM plates for the ability to grow on methanol. Subsequently, the complementing plasmids were rescued from the positive clones in *E. coli* DH5 $\alpha$ . To facilitate sequencing, a complementing 1.6-kb *Eco*RI DNA fragment was subcloned into the *Eco*RI-digested phagemid pBluescript II SK<sup>+</sup> (Stratagene), resulting in pBSK-PEX12. Sequencing of both strands was carried out on a LiCor automated DNA-sequencer using dye-primer chemistry. For DNA and amino acid sequence analysis, the Clone Manager 5 program (Scientific and Educational Software, Durham, N.C.) was used. The TBLASTN algorithm (Altschul et al. 1997) was used to search the databases at the National Center for Biotechnology Information (NCBI, Bethesda, Md.). The nucleotide sequence of *H. polymorpha* *PEX12* was deposited in Genbank and was assigned the Accession No. AF333026.

### Mutant isolation and analysis

*H. polymorpha* HF246 cells were transformed with *Bam*HI-linearised pREMI-Z DNA in the presence and absence of *Bam*HI (1 U/ $\mu$ g). Transformants were selected by incubating on YPD-zeocin plates for 2 days at 37°C, replica-plated to agar plates with mineral medium containing 0.5% (v/v) methanol as the sole carbon source and incubated for 2–3 days at 37°C. Methanol non-utilising (Mut<sup>-</sup>) strains were selected. For ultrastructural analysis, Mut<sup>-</sup> mutants from HF246 were induced on methanol-containing plates or grown in glycerol/methanol mixtures. Peroxisomal profiles were analysed by fluorescence microscopy as described by Baerends et al. (2000). Mutants involved in selective peroxisome degradation were

isolated using the same strain according to the method of Titorenko et al. (1995), as well as by screening for prolonged eGFP-SKL fluorescence after a shift to conditions in which peroxisomes are normally degraded. Total DNA was isolated from selected transformants and approximately 5 µg of total DNA was digested with *EcoRI*, which does not cleave pREMI-Z. After complete digestion the DNA was self-ligated and used to transform *E. coli*. Plasmid DNA was recovered from zeocin-resistant colonies. pREMI-Z plasmids containing genomic inserts were directly sequenced using the adapter primers or the primers pREMI-ori (5'-GATCTTTTCTACGGGGTCTG-3') and pREMI-tef (5'-CG-GAGTCCGAGAAAATCTGG-3').

## Results

### Restriction enzyme-mediated integration events in *H. polymorpha*

Classical mutagenesis by NTG or UV treatment has been the routine tool for isolation of specific mutants in the yeast *H. polymorpha*. However, these screens, and the isolation of the corresponding genes by functional complementation, are laborious and time consuming. It would therefore be highly desirable to develop an insertional mutagenesis procedure, like REMI, which is more efficient than the methods used to date. Initial experiments in *H. polymorpha* using plasmids containing homologous auxotrophic markers were unsuccessful due to the high frequency of gene conversion events (data not shown). Previously, we demonstrated that the zeocin resistance gene from *Streptoalloteichus hindustanus* (Drocourt et al. 1990), under the control of the *S. cerevisiae* *TEF1* promoter, could be used as a dominant selection marker in *H. polymorpha* (Salomons et al. 2000). Thus, we constructed a plasmid, designated pREMI-Z, that does not contain any autonomously replicating sequences for *H. polymorpha* (see Fig. 1A). Its main features are (1) the zeocin resistance gene for selection in *H. polymorpha* and *E. coli* (Gatignol et al. 1988); (2) central *Bam*HI, *Tth*111I and *Nhe*I sites, which can be used for insertional mutagenesis in *H. polymorpha*; and (3) binding sites for the M13/pUC reverse and universal sequencing primers next to the central restriction sites that allow direct sequence analysis of the chromosomal DNA fragments flanking the integrated plasmid without the need for subcloning. Its small size (approximately 2.0 kb) and the fact that it has few restriction sites allows the use of many different restriction endonucleases for the recovery of genomic fragments.

To test whether pREMI-Z functions in insertional mutagenesis by REMI, we transformed wild-type *H. polymorpha* cells with circular plasmid DNA, *Bam*HI-linearised DNA, and *Bam*HI-linearised DNA supplemented with increasing amounts of the restriction enzyme *Bam*HI (Fig. 2). Surprisingly, large numbers of transformants were obtained with *Bam*HI-linearised pREMI-Z DNA. Addition of the restriction enzyme *Bam*HI to the transformation mixture enhanced the transformation frequency. However, the stimulation was

only moderate, reaching a maximum of 1.8-fold when 1 U of restriction enzyme was added (see Fig. 2). Addition of larger amounts of restriction enzyme negatively influenced the transformation frequency. These data suggest that the bulk of transformants do not result from REMI events.

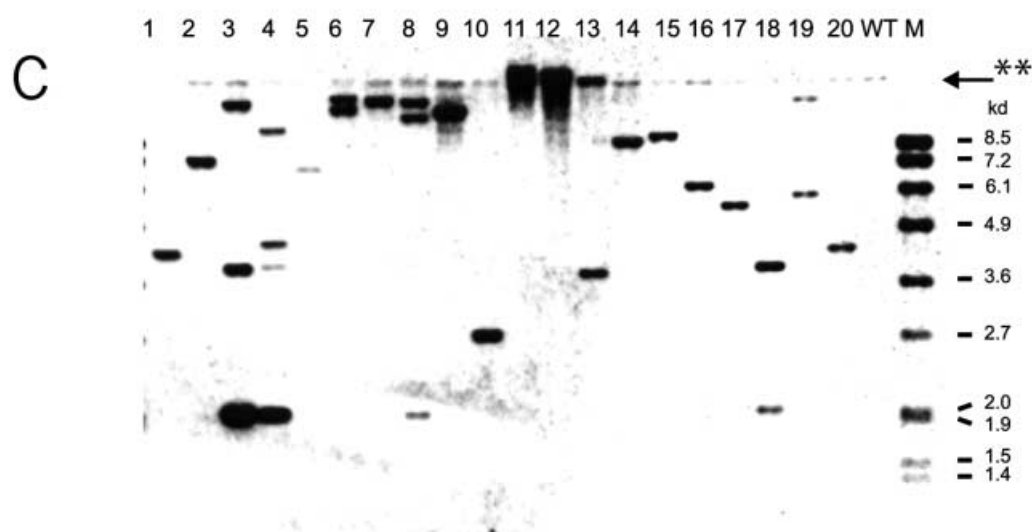
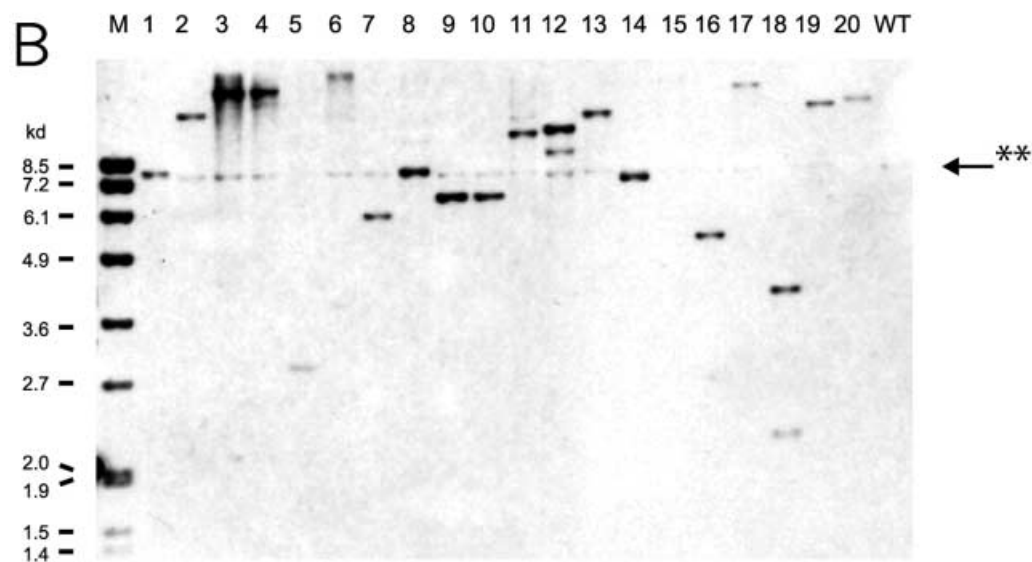
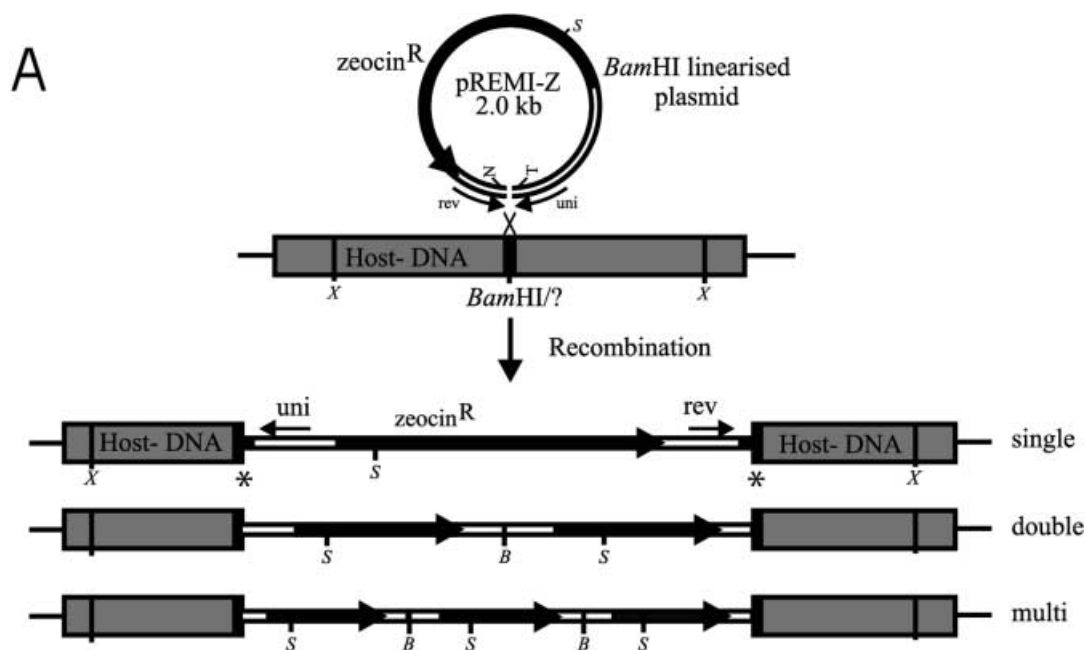
pREMI-Z integrates stably and randomly into the *H. polymorpha* genome

To test the stability of pREMI-Z transformants we cultivated 10 randomly picked transformants for several generations (~30) in non-selective medium (YPD). In all cases no loss of zeocin resistance was observed, indicating stable integration of the pREMI-Z cassette into the *H. polymorpha* genome.

In order to investigate whether the pREMI-Z cassette had integrated at random or site-specifically into the genomic DNA of the host, we analysed 40 randomly chosen transformants – obtained by transformation with linear pREMI-Z in the presence of *Bam*HI – by Southern hybridisation. Total genomic DNA from these transformants was digested with *EcoRI*, and subjected to Southern analysis using pREMI-Z as a probe. The restriction enzyme *EcoRI* does not digest pREMI-Z (Fig. 1A) and should therefore generate a single hybridising fragment of variable size when the linearised plasmid integrates at random in a single genomic locus. Indeed, in almost all cases one hybridising fragment, ranging from 2.1 kb up to at least 15 kb in size, was observed (shown for 20 transformants in Fig. 1B). However, in two of the transformants two hybridising fragments (Fig. 1B, lanes 12 and 18) were observed, indicating that the pREMI-Z cassette might have integrated at two different genomic loci. These data indicate that no clear integration hotspots are present in the *H. polymorpha* genome, but that pREMI-Z apparently integrates in a random fashion.

It has been shown for many organisms that the inclusion of restriction endonuclease during transforma-

**Fig. 1A–C** Integration of pREMI-Z into the genome of *H. polymorpha*. **A** Schematic model of the integration of the linear pREMI-Z cassette into the *H. polymorpha* genome. Single, double and multicopy integrations are depicted and include the regeneration of the internal *Bam*HI sites. Restriction enzymes that do not digest pREMI-Z (X; in this study, *EcoRI*) can be used to isolate the pREMI-Z cassette together with flanking genomic DNA fragments. The *Sma*I site in pREMI-Z can be used to remove additional copies of the vector from rescued plasmids. Abbreviations: N, *Nhe*I; T, *Tth*111I; B, *Bam*HI; S, *Sma*I; rev, M13/pUC reverse primer; uni, M13/pUC universal primer. The asterisks indicate the site of integration. **B**, **C** Representative Southern analyses of 20 randomly picked transformants obtained using *Bam*HI-linearised pREMI-Z in the presence of *Bam*HI in the transformation mixture. Genomic DNA was digested with *EcoRI* (**B**) or *Bam*HI (**C**). In both cases *Bam*HI-linearised pREMI-Z was used as probe. Bands marked \*\* are non-specific hybridisation signals. Fragment lengths of the molecular weight marker (bacteriophage *SPPI* DNA digested with *EcoRI*) are given in kb



tion directs the insertion of exogenous DNA to corresponding cleavage sites in the host genome (Schiestl and Petes 1991). To test whether pREMI-Z integrated preferentially at *Bam*HI sites in *H. polymorpha*, we digested the chromosomal DNA of the 40 selected transformants with *Bam*HI and performed a Southern analysis. When *Bam*HI-linearised pREMI-Z DNA integrates at *Bam*HI sites in the host genome and these restriction sites are restored, this analysis should detect a single hybridising band of the same size as pREMI-Z. In none of the transformants was this the case (shown for 20 transformants in Fig. 1C). In contrast, we observed either single hybridising bands of sizes larger than pREMI-Z, or multiple hybridising bands of different sizes, sometimes in combination with a fragment the size of pREMI-Z (Fig. 1C). This suggests that, even in the presence of restriction endonuclease, *Bam*HI sites were not the major targets for integration of the plasmid. Alternatively, it is possible that the restriction sites were not regenerated at both ends during integration.

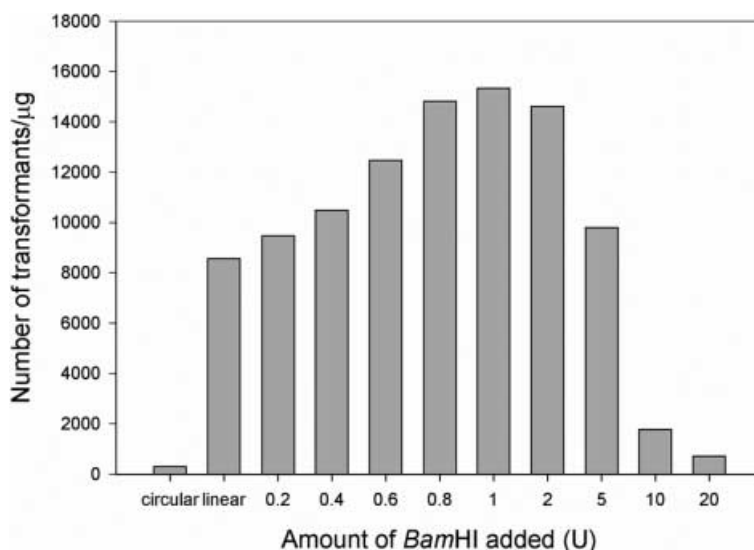
In addition, Southern hybridisation analysis was used to determine the copy number of the pREMI-Z cassette in the transformants, based on the number of hybridisation signals observed. This analysis revealed that in 50% of the transformants a single molecule of the cassette was present in the *H. polymorpha* genome. Two copies of the pREMI-Z cassette were found in 25% of the transformants analysed (e.g. Fig. 1C, lanes 6, 13 and 19). Three or more copies of the cassette were observed in the remaining (25%) cases (e.g. Fig. 1C, lanes 3, 4 and 8). Thus, in strains harbouring multiple copies of pREMI-Z, the internal *Bam*HI sites on both sides of the cassette were apparently restored, resulting in three hybridisation signals, including one the size of pREMI-Z (schematically represented in Fig. 1A; see Fig. 1C, lanes 3 and 4). The fact that in these multicopy situations the internal *Bam*HI restriction site(s) can be cleaved indicates that the tandem arrangement of the cassette had probably occurred prior to integration. Such tandem

integration was also observed in some pREMI-Z transformants obtained in the absence of *Bam*HI in the transformation mixture (data not shown).

#### Generation of *H. polymorpha* mutants using pREMI-Z

An initial screen of a limited number of *H. polymorpha* pREMI-Z transformants revealed that it was possible to create leucine, uracil and adenine auxotrophic mutants by gene tagging. To investigate the capacity of pREMI-Z to tag genes of interest in *H. polymorpha* in more detail, a large-scale mutagenesis was performed. We chose to isolate mutants defective in the biogenesis and in the selective degradation of peroxisomes on the basis of our ongoing research in these fields (Veenhuis et al. 2000). Peroxisomes develop in response to specific metabolic needs. In *H. polymorpha*, peroxisomes proliferate and become abundant when methanol is used as sole carbon and energy source. Conversely, the organelles are specifically degraded via a process designated pexophagy when they become redundant for growth (Veenhuis et al. 1983). The crucial enzymes required for methanol metabolism are located in peroxisomes, and mutants defective in peroxisome biogenesis (*pex* mutants) display a specific growth defect on methanol-containing medium (Mut<sup>-</sup>) (van der Klei et al. 1991). This screening method is well established for *H. polymorpha* and has contributed to a large extent to the identification of *PEX* genes from this yeast. Our screen for mutants disturbed in selective peroxisome degradation (*pdd* mutants) is based on the failure of peroxisomes to disappear after shifting cells from peroxisome-proliferating conditions to conditions not requiring peroxisomes for growth (Titorenko et al. 1995). To allow efficient screening of mutants we constructed a peroxisomal import reporter strain, designated HF246. This strain produces the eGFP protein, C-terminally extended with a peroxisomal targeting

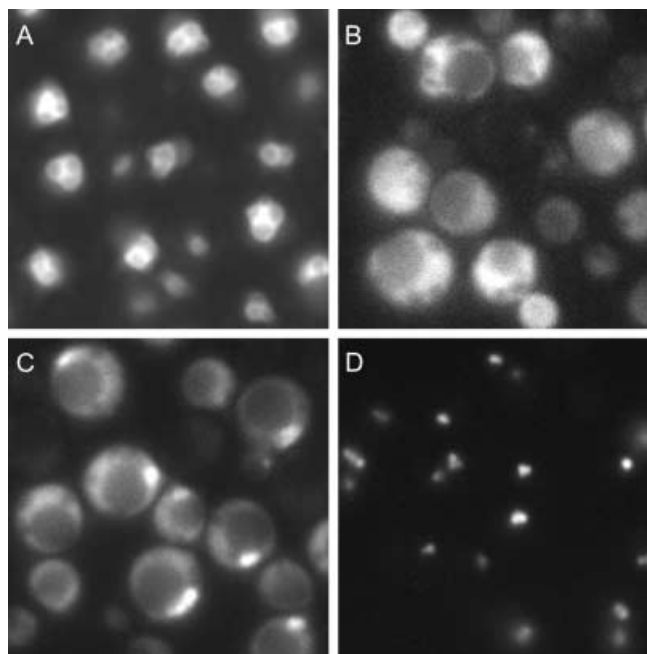
**Fig. 2** Transformation frequencies of pREMI-Z in *H. polymorpha*. A 1- $\mu$ g aliquot of circular or *Bam*HI-linearised pREMI-Z was used to transform *H. polymorpha* NCYC495 (*leu1.1*) or HF246 (*leu1.1*) cells. The restriction enzyme *Bam*HI was added in increasing amounts prior to transformation. The results are the averages of three independent experiments



signal, -serine-lysine-leucine-COOH (SKL), under the control of the strong, inducible, alcohol oxidase promoter. In wild-type cells, eGFP-SKL is efficiently targeted to peroxisomes and does not interfere with normal growth on methanol. Under these conditions eGFP-SKL is present at the rim of the peroxisomes (Fig. 3A), because the major peroxisomal matrix proteins, alcohol oxidase and dihydroxyacetone synthase, are present in the peroxisomal matrix in a crystalloid structure. This strain allowed us to monitor the distribution of eGFP-SKL throughout the cells, and could therefore be used to discriminate mutants with altered peroxisomal profiles (eGFP-SKL mislocalisation, aberrant number and size of peroxisomes) amongst *Mut<sup>-</sup>* transformants. In addition, the prolonged presence of eGFP-SKL fluorescence after a shift to conditions that should induce selective peroxisome degradation provided an additional screen for the isolation of *pdd* mutants.

We transformed the *H. polymorpha* reporter strain HF246 with *Bam*HI-linearised pREMI-Z in the presence and absence of *Bam*HI endonuclease as described above. After selection on YPD-zeocin medium, the transformants – approximately 5500 and 4500 obtained with and without *Bam*HI, respectively – were replica plated onto methanol-containing medium. Mutants affected in growth on methanol were selected for further analysis. Approximately equal amounts of *Mut<sup>-</sup>* mutants were obtained in both the presence and the absence of added restriction endonuclease. An initial fluorescence analysis of 195 methanol-induced *Mut<sup>-</sup>* mutants (from plates) revealed 72 potential candidates with aberrant peroxisomal structures; i.e. with eGFP-SKL profiles differing from that of the wild type. The other strains displayed normal peroxisomes with peroxisomal eGFP-SKL, although they had been isolated as *Mut<sup>-</sup>* mutants. These mutants are most probably affected in peroxisomal enzymes or in steps in methanol metabolism unrelated to peroxisome structure. Subsequently, the mutants that displayed an aberrant eGFP-SKL pattern/staining/distribution were grown in glycerol/methanol-containing medium and analysed by fluorescence microscopy in more detail. Distinct groups of mutants could be recognised (exemplified by Fig. 3B–D). The first class, *pex* mutants, displayed a distinct cytosolic distribution pattern of eGFP-SKL (Fig. 3B) compared to the normal fluorescent ring staining observed in wild-type cells (compare Fig. 3A and B). In these mutants peroxisomes are not present at all, or import of eGFP-SKL is completely blocked. The second group of mutants displayed a phenotype previously described as *Pim<sup>-</sup>* (Waterham et al. 1992). This phenotype is characterised by a partial defect in the import of matrix proteins into peroxisomes, as indicated by the presence of both cytosolic and peroxisomal eGFP-SKL (Fig. 3C). Other groups of mutants appeared to be affected in peroxisome morphology, but not in the import of the eGFP-SKL marker protein (see Fig. 3D).

In addition, using the same conditions, we were able to isolate mutants affected in selective peroxisome de-



**Fig. 3A–D** Localisation of eGFP.SKL in *H. polymorpha* HF246 *Mut<sup>-</sup>* mutants obtained by gene-tagging. In wild-type *H. polymorpha* cells a typical staining of the peroxisome rim can be observed (A), while in *pex* mutants eGFP-SKL remains in the cytosol (B). In *pim* mutants (C) small labelled peroxisomal structures can be identified, as well as cytosolic eGFP-SKL. An example of altered peroxisomal morphology is shown in D, here exemplified by the presence of very small peroxisomes. Cells were induced in methanol/glycerol-containing medium at 37°C

gradation (24 *pdd* mutants were isolated from approximately 4000 pREMI-Z transformants). These mutants displayed a reduced ability to degrade peroxisomes as revealed by reduced degradation of the peroxisomal matrix protein alcohol oxidase as well as the prolonged presence of eGFP-SKL fluorescence after shifting cells to conditions not requiring functional peroxisomes (data not shown). Taken together, these results demonstrate that gene tagging is a highly efficient method for generating mutants in *H. polymorpha*.

#### pREMI-Z integration is responsible for the phenotype of the isolated mutants

In order to determine whether rescue of the pREMI-Z cassette with genomic flanking regions was possible, and also to obtain detailed information as to how pREMI-Z had integrated into the *H. polymorpha* genome, we prepared genomic DNA from 36 *Mut<sup>-</sup>* mutants from the two mutant collections (i.e. with and without *Bam*HI in the transformation mixture) that displayed an aberrant eGFP-SKL localisation. The *pdd* mutants were not included at this stage since, so far, only limited sequence information is available for *PDD* genes. As an initial choice of restriction enzyme, we tested if the integrated plasmids, together with the flanking genomic sequences, could be rescued after digestion of the genomic DNA

with *EcoRI*, although in principle any restriction enzyme that does not digest pREMI-Z can be used for this purpose. From most of the mutants, we were able to obtain the pREMI-Z vector together with flanking genomic fragments after transformation of the self-ligated *EcoRI*-digested genomic DNA into *E. coli*. In cases where we failed to obtain transformants, the flanking genomic fragments obtained by digestion with *EcoRI* were probably too large to be rescued in *E. coli* (see e.g. lanes 6 and 17 in Fig. 1B). However, in such cases we were able to rescue the pREMI-Z cassette using other restriction enzymes to digest the genomic DNA. We subsequently tested whether the rescued plasmids carrying a genomic insert harboured multiple pREMI-Z cassettes as found in the Southern analysis (Fig. 1C). For this purpose, we digested the plasmid DNA with *SmaI*, which will give rise to a DNA fragment the size of pREMI-Z when multiple cassettes are present in the rescued plasmid (Fig. 1A). In principle, any restriction enzyme that cleaves once in the plasmid and not in the insert can be used to eliminate the additional copies of the cassette. Thirteen of the 36 transformants contained such a multicopy arrangement of the pREMI-Z cassette (these were obtained from both mutant collections). To allow sequencing of the inserts in these plasmids, the additional copies were removed by self-ligation of the isolated *SmaI* fragment. Subsequently, all rescued plasmids were amplified and sequenced.

Analysis of the sequences obtained (Table 1 and Fig. 4) revealed that in 17 out of 36 selected mutants the cassette had integrated in or near genes known to be involved in peroxisome biogenesis/function in *H. polymorpha*. Thus the Mut<sup>-</sup> phenotypes obtained in these cases are clearly linked to the integration of the pREMI-Z cassette into the *H. polymorpha* genome. In addition,

we identified nine new genes for which the function in peroxisome biogenesis is as yet unknown. Finally, sequences from 10 mutants did not reveal any clue to which gene was responsible for the observed Mut<sup>-</sup> phenotype and the aberrant eGFP-SKL localisation. In these cases the nucleotide sequences either did not display any similarity to sequences available in the current databases, or the regions flanking pREMI-Z showed similarity to two different genes. In rare cases, integration of the pREMI-Z cassette had occurred in the promoter region of two divergently transcribed genes and the observed phenotype could therefore not be directly linked with either of the two genes.

A detailed comparison of the sequences obtained with sequences of known *H. polymorpha* genes and sequence tags (STS; Blandin et al. 2000) showed that in the mutant collection obtained in the absence of *BamHI*, integration of pREMI-Z had never occurred at *BamHI* sites (Fig. 4). Instead, the integration appeared to occur at random, since no sequence similarity was observed between the different integration sites. In contrast, in the mutants from the collection obtained in the presence of *BamHI*, approximately 30% of the integrations had occurred at a *BamHI* site. However, in most of these cases, the *BamHI* site was not regenerated. Rather, we frequently observed that one or a few nucleotides had been lost from pREMI-Z during the integration of the cassette (Fig. 4).

As noted above, we sometimes observed that the sequences on the two flanks of pREMI-Z were not part of the same gene, but rather showed identity to two different *H. polymorpha* sequences (Fig. 4, mutants RVD3, MVH6, KNF4), suggesting the occurrence of deletions in the host genome. One such recovered plasmid contained a DNA fragment with similarity to the *P. pastoris* *PEX12*

**Table 1** Analysis of selected *H. polymorpha* mutants affected in peroxisome biogenesis and/or function obtained by gene-tagging

| Gene class                                    | Genes identified <sup>a</sup>  |  |
|---|--|--|
|   | Peroxisome biogenesis  | Growth on methanol                                       |
| Genes with a known function<br>(17 mutants)   | <i>PEX1</i> *<br><i>PEX5/PEX5</i> *<br><i>PEX6/PEX6</i> *(2)<br><i>PEX8</i> (2)<br><i>PEX12</i> *<br><i>PEX13</i><br><i>PEX14</i> *(2) | <i>CAT</i><br><i>DAS</i><br><i>DAK</i> (2)/ <i>DAK</i> * |
| Genes with an unknown function<br>(9 mutants) | Peroxisome biogenesis<br>relation unknown<br>Hypothetical proteins<br>putative transcription factors                                   |  |
| Unclear integration events<br>(10 mutants)    | No similarity<br>Different similarity on two<br>flanking regions<br>In promoter region of<br>divergent genes                           |  |

<sup>a</sup>Sequence analysis of 36 selected Mut<sup>-</sup> mutants displaying aberrant eGFP-SKL fluorescence patterns, followed by database searches at the NCBI, resulted in the identification of genes tagged by pREMI-Z. The identified genes are listed with respect to their possible relationship to peroxisome biogenesis/growth on methanol. The asterisks indicate cases where the sequence similarity with the indicated gene was observed in only one region flanking pREMI-Z in the rescued plasmid. In these cases, the other flanking sequence showed no similarity or showed similarity to a different gene. The numbers of independent cases isolated are given in parentheses



gene on only one flanking region (mutant RVD3). *PpPex12p* encodes an integral peroxisomal membrane protein with a zinc-finger motif, and is involved in peroxisome biogenesis (Kalish et al. 1996). Since this gene had not previously been isolated from *H. polymorpha*, and the integration of the pREMI-Z cassette had not occurred via a simple integration event, we analysed this mutant in more detail. Initial sequence analysis of the RVD3 mutant showed that the cassette had integrated at a possible *Bam*HI site at the 5' end of the *PEX12* gene. However, the genomic sequence from the other flank of the rescued insert did not show any similarity to the previously identified *P. pastoris PEX12* gene, but rather to an uncharacterized STS, suggesting a possible deletion in the genome of the mutant (Fig. 4). To obtain further information about the integration event and the possible deletion, we functionally complemented the RVD3 mutant using a genomic *H. polymorpha* library (Tan et al. 1995). A 1.6-kb complementing *Eco*RI fragment was isolated and sequenced. Sequence analysis revealed that the fragment indeed contained an ORF predicted to encode a protein that was highly similar to *P. pastoris Pex12p* over the entire length of the protein (51% identity, data not shown). Therefore, we designated this gene *H. polymorpha PEX12* (*HpPEX12*). Analysis of the DNA sequence of *HpPEX12* showed that pREMI-Z had integrated between two *Bam*HI restriction sites in the *H. polymorpha* genome, resulting in the deletion of (a) *Bam*HI fragment(s) of unknown size. In RVD3 the promoter region and the first 52 codons of the *HpPEX12* gene have been deleted, which most probably caused the observed *Mut*<sup>-</sup> phenotype and the aberrant eGFP-SKL localisation in the RVD3 mutant. However, we cannot exclude the possibility that deletion of other genes upstream of *HpPEX12* also contributes to the observed phenotype. Further analyses of the role of *PEX12* in peroxisome biogenesis/matrix import are currently underway. However, these are beyond the scope of the present study.

## Discussion

In the present study we show that insertional mutagenesis is a powerful tool for tagging genes of interest in the methylotrophic yeast *H. polymorpha*. We have evaluated the use of restriction enzyme-mediated integration (REMI) in *H. polymorpha* using the non-replicating plasmid pREMI-Z. Surprisingly we observed that a high frequency of random integration occurred in the absence of added restriction enzyme. Since this process is fundamentally different from REMI we designated this method Random integration of Linear DNA Fragments (RALF).

In most microorganisms tested, it has been observed that in REMI the addition of restriction enzyme prior to transformation dramatically enhances the transformation frequency (Granado et al. 1997; Brown et al. 1998; Sanchez et al. 1998). Most probably, this occurs because cycles of digestion and ligation can occur, whereby linear plasmid DNA can be integrated into

cleavage sites in the chromosomes. Indeed, addition of restriction enzyme to linear pREMI-Z increased the transformation frequency of *H. polymorpha* cells somewhat. However, the relative increase in transformation frequency (only 1.8-fold) is much lower than the values reported for other species (up to 60-fold) (Brown et al. 1998; Sanchez et al. 1998). Apparently, the *H. polymorpha* transformation system can process linear DNA much more efficiently than the transformation systems used by other organisms in which REMI was tested ( $\approx 100$ –1000 times more transformants/ $\mu$ g DNA), a feature that has been observed before (Faber et al. 1994b). Also the amount of restriction enzyme (0.6–5 U) that resulted in significant numbers of *H. polymorpha* transformants was much lower than the 5–250 U that were used in other organisms to improve the transformation frequency (Kuspa and Loomis 1992; Sato et al. 1998). We observed that 10 U of restriction enzyme negatively affected the number of transformants, most probably by causing double-strand breaks in the genomic DNA which are inefficiently repaired.

Our data also indicate that linearised pREMI-Z integrated randomly into the genome rather than at restriction sites. Thus, in the absence of restriction endonuclease, integration of linear pREMI-Z DNA in the *H. polymorpha* genome does not occur via REMI, but rather via a process we designated RALF. Addition of restriction enzyme resulted in the appearance of an additional population of mutants, characterised by integration of pREMI-Z at restriction sites, suggesting that these represent true REMI events. However, since the number of these additional mutants was rather limited, addition of restriction endonuclease is not required for mutagenesis. It should be possible in *H. polymorpha* to tag via RALF any gene that it is not essential for cell viability. This is in clear contrast to REMI methods reported for other yeasts, like *C. albicans* (Brown et al. 1996) and *S. cerevisiae* (Schiestl and Petes 1991), where integration is critically dependent on the addition of restriction enzyme and integration occurs almost exclusively at the corresponding recognition sites, thus limiting the use of REMI for random gene tagging. In contrast, the high frequency of illegitimate recombination in *H. polymorpha* allows genome-wide mutagenesis by this method.

We have observed that, in almost 50% of the transformants analysed, multicopy integration of the pREMI-Z cassette had occurred at a single genomic locus. These tandem integrations have also been observed in other organisms (Granado et al. 1997; Sato et al. 1998). It seems likely that the pREMI-Z cassette forms multimers prior to integration, a process that is independent of the addition of restriction enzyme. One possible explanation for these multicopy arrangements is that in some cases zeocin resistance requires a higher level of transcription of the zeocin resistance cassette than can be obtained with a single copy of pREMI-Z. This might depend on the site of integration in the genome. Like the

|            |  |
|------------|--|
| pREMI-Z    | AAACGACGGCCAGTCTAGCTAGCTA <b>GGATCC</b> TGACTGACTGACATGGTCATAGCTG                                |
| RVD3-ori   | CTAGCTA-GATCCAAATAACAAAGCTCCCTTGTGCGGGCCGCGAMGCTCTAGTA   |
| Hp-STS     | ACAACCACGAACCTTACCAGGAAATTGCCATTGGCGAGCT <b>GGATCC</b> AAATAACAAAGCTCCCTTGTGCGGGCCGCGA-CCTCTAGTA |
| RVD3-tef   | AGTCCTCTGATTGCAAGATTCAAGTTCATCGAAGTGGTTA <b>GGATCC</b> TGACTGACTGACATG                           |
| Hp-PEX12   | AGTCCTCTGATTGCAAGATTCAAGTTCATCGAAGTGGTTA <b>GGATCC</b> TGATAAGATAGCGGGGTATCTATTGGCGTAGTGGACCAG   |
| ADN-P3-ori | CAGTCTAG-----GCTCTCCAAGACAGCCTGGTCTTGTGCATTACCTATCAGG  |
| Hp-PEX8    | TGTAGGTATCGACGCTGGTGAAGTATTAAAGCATGACTGAATGTGCGCTCTCCAAGACAGCCTGGTCTTGTGCATTACCTATCAGG           |
| ADN-3-tef  | TGTAGGTATCGACGCTGGTGAAGTATTAAAGCATGACTGAATGTG-----ACATG  |
| ADN-T6-ori | CAGTCTAGCTAGCTAGGACAGGCGCTTTGAGGTGCGCATGGACCAGGTCTTGCTTGTAAAT                                    |
| Hp-DAK     | CACACGCTCGGATTGATGAATTGGAGGTCTGGATTGGCATGACACAGGCTTTGAGGTGCGCATGGACCAGGTCTTGCTTGTAAAT            |
| ADN-T6-tef | CACACGCTCGGATTGATGAATTGGAGGTCTGGATTGGCATG-TCCTGACTGACTGACATG                                     |
| MVH6       | CAGTCTAGCTAGCTAGGNCCTAAAGCTAAATCACACCGCTGGTITTTTTTCTTGCAGCTGC                                    |
| Hp-STS     | GCAACCCGTCTCCCATGTAAACGTACTAACAGACAAAACGCACTAAAGCTAAATCACACCGCTGGTITTTTTTCTTGCAGCTGC             |
| MVH6-tef   | AGATCACCACTAAAACAGCCAAGTTTCTGTAAACCCGAAGAGCNCCTGACTGACTGACATG                                    |
| Hp-PEX6    | AGATCACCACTAAAACAGCCAAGTTTCTGTAAACCCGAAGAGCATACGAGGCAAAAAGCTGATCATCTCATGATCAAATTCGTG             |
| MVH10-ori  | CAGTCTAGCTA---GATCCTAATCAGCAACAGAAAGATTCTATGTCGCTAAGGCTTGCT                                      |
| Hp-PEX6    | TACTTTATATCCGGCACATTGAAGCACTTGCAAAAAGCC <b>GGATCC</b> TATCAGCAACAGAAAGATTCTATGTCGCTAAGGCTTGCT    |
| MVH10-tef  | TACTTTATATCCGGCACATTGAAGCACTTGCAAAAAGCC <b>GGATC</b> -----GACTGACATG                             |
| ARJ8-ori   | CAGTCTAGCTAGCTAGACATATTTGGCAGGTTGTTGCAGACGGTCGAGAGTAGCAAGCAGC                                    |
| Hp-STS     | CAAAGGCGACAGAAAGCCTGGGCTCCTAGAGCGATGGAGTACATATTTGGCAGGTTGTYGCAGACGGTCGAGAG-AGCAAGCAGC            |
| ARJ8-tef   | CAAAGGCGACAGAAAGCCTGGGCTCCTAGAGCGATGGAGT-TCCTGACTGACTGACATG                                      |
| ARJ11-ori  | CAGTCTAGCTAGCTAGGACGCTGCAGCGTTTGGTTCGCGATTCTCTTTTAAAGGAATAGTGT                                   |
| Hp-STS     | TCTAGTCACTATGGCTGAAGCGGTTAATCAGTAAATGATGTCCGCTGCAGCGTTTGGTTCGCGATTCTCTTTTAAAGGAATAGTGT           |
| ARJ11-tef  | TCTAGTCACTATGGCTGAAGCGGTTAATCAGTAAATGATGTG-CTGACTGACTGACATG                                      |
| ARJ59-ori  | CAGTCTAGCTAGCTAGGATTCTAGTGGTCTAAGCGTGAGCCTCCCTCTCTGACCAACAA                                      |
| Hp-DAS     | CACGCATGCGCCCCCTCTCCATCTCCTCTCGACAGCGATCTGTTTTCTAGTGGTCTAAGCGTGAGCCTCCCTCTCTGACCAACAA            |
| ARJ59-tef  | CACGCATGCGCCCCCTCTCCATCTCCTCTCGACAGCGATCTG-CTGACTGACTGACATG                                      |
| ARJ55-ori  | CAGTCTAGCTAGCT---CCTTATATGGTGGTGGGATGTATGGTAATCGCTACGGAGGTA                                      |
| Hp-PEX13   | CATGTACAACCTTTACGGATCATATGGATCGTCAATGTAC <b>GGATCC</b> TTATATGGTGGTGGGATGTATGGTAATCGCTACGGAGGTA  |
| ARJ55-tef  | CATGTACAACCTTTACGGATCATATGGATCGTCAATGTAC-GATCCTGACTGACTGACATG                                    |
| KNF1-ori   | CAGTCTAGCTAGCTAGGACAAACGAGCGTTTGCTACGCTGGAGAGATGGATCGAAACCAA                                     |
| Hp-PEX5    | CACCTATGACTCTGGCAATTCGTACATTAAACGAAGGGTACGACAAACGAGCGTTTGCTACGCTGGAGAGATGGATCGAAACCAA            |
| KNF1-tef   | CACCTATGACTCTGGCAATTCGTACATTAAACGAAGGGTACG--CTGACTGACTGACATG                                     |
| KNF16-ori  | CAGTCTAGCTAGCTAGGAGCAAAAGCTGCCAGAATGGAAGGATCCAAGACCTTTTCTAAAG                                    |
| Hp-STS     | TACGTGGTTTTCAGAAGGCTGGTATTCAGTITTTTGGCCCTAGTGCAAAAGCTGCCAGAATGGAAGGATCCAAGACCTTTTCTAAAG          |
| KNF16-tef  | TACGTGGTTTTCAGAAGGCTGGTATTCAGTITTTTGGCCCTAG-TCCTGACTGACTGACATG                                   |
| KNF19-ori  | CAGTCTAGCTAGCT--CAACAACCTATAGGTTTACGCTTAAAAATGCAGATTTTCGTAAAG                                    |
| Hp-STS     | AGCAGCATATAAATATCGCGCATCCCCTAACTTTTAAAGACAACACCTATAGGTTTACGCTTAAAAATGCAGATTTTCGTAAAG             |
| KNF19-tef  | AGCAGCATATAAATATCGCGCATCCCCTAACTTTTAA--GATCCTGACTGACTGACATG                                      |
| KNF33-ori  | CAGTCTAGCTAGCTAGGACACCACTCGCCCTTGGAATTGTACCATTTGTAGGTGTGGCCA                                     |
| Hp-CAT     | GATTGTGCACGCCCTTGGTTGGCAATGAAGTGGACCTGCACATACCCACTCGCCCTTGGAATTGTACCATTTGTAGGTGTGGCCA            |
| KNF33-tef  | GATTGTGCACGCCCTTGGTTGGCAATGAAGTGGACCTGCACA--TCCTGACTGACTGACATG                                   |
| KNF38-ori  | CAGTCTAGCTAGC--TACACGGGCGGTACACCACGTCGCTAGACGGTCTCGTGGTITTTTCCA                                  |
| Hp-DAK     | GCAGCTGGCTACCGACTACAAGATCAAGCCCGCAAGAGTGTACACGGGCGGTACACCACGTCGCTAGACGGTCTCGTGGTITTTTCCA         |
| KNF38-tef  | GCAGCTGGCTACCGACTACAAGATCAAGCCCGCAAGAGTG--TCCTGACTGACTGACATG                                     |
| KNF47-ori  | CAGTCTAGCTAGCTAGGACTCACTGACAGGAAAATTTTAAAGCGCGCACCAAGTGGCTGG                                     |
| Hp-STS     | GTTGCAAAATTTAAAAATTTTTCAGTCATATTCACTTCCATCCAGCTCACACTGACAGGAAAATTTTAAAGCGCGCACCAAGTGGCTGG        |
| KNF47-tef  | GTTGCAAAATTTAAAAATTTTTCAGTCATATTCACTTCCATC-CA-TCCTGACTGACTGACATG                                 |
| KNF4-ori   | CAGTCTAGCTAGCTAGACTCTTCAAAATTCGATAAAATACTACTTTTCCATCTGCAATCAT                                    |
| Hp-STS     | AAAATTTCTTTGAAATCGCGTAATCCCTGCATTGACAGTGCCTCTTCAAAATTCGATAAAATACTACTTTTCCATCTGCAATCAT            |
| KNF4-tef   | CAAGCTCCAGGGCGAGCCAGAACCCACATTTTGGAGCTG <b>TATGCN</b> NGNCTGNCTGACA-G                            |
| Hp-STS     | MAAGCTCCAGGGTGAGCTCAAAAGCYACATTTTGGAGCTGATGGTAAGGAACAGTACCAATTTCTGGTGACGCGCCTTTAAGGGCT           |
| KNF28-ori  | CAGTCTAGCTAGCTAGGATCCACTTTCTTGAATGCTACGCTGCTGGTCCAGCATTTGATGA                                    |
| Hp-PEX8    | TTGCAGATCAAAACAGCAGTTTCCCGGCTCTGTAGCAGAATTTGTCACATTTCTTGAATGCTACGCTGCTGGTCCAGCATTTGATGA          |
| KNF28-tef  | TTGCAGATCAAAACAGCAGTTTCCCGGCTCTGTAGCANNAT <b>TTGTC</b> CTGACTGNCTGACA-G                          |
| KNF8-ori   | CAGTCTAGCTAGCTA--GCAGGGCCCTCTGCAAGCAGCATCGCCAACCATGCAGTACAC                                      |
| Hp-DAS     | TTGTCCAGCCCATGTGGCCGGCCAGCGAGATCGACTCGAGAGCAGGGCCCTCTGCAAGCAGCATCGCCAACCATGCAGTACAC              |
| KNF8-tef   | TTGTCCAGCCCATGTGGCCGGCCAGCGANATAGACTCGAGA-CTGANTGNCTGACATG                                       |

**Fig. 4** Sequence analysis of selected Mut<sup>-</sup> mutants obtained by gene tagging in the presence and absence of restriction endonuclease. pREMI-Z plasmids with their genomic flanking regions were isolated from selected Mut<sup>-</sup> mutants with altered peroxisomal profiles and sequenced using the pREMI-ori (ori) and pREMI-tef (tef) primers. Nucleotide sequences were aligned with known *H. polymorpha* genes and sequence tags (STS) from databases at NCBI. The upper panel shows sequences from mutants obtained in the presence of BamHI in the transformation mixture, the lower panel those obtained in the absence of the enzyme. Vector sequences are shown in *italics*. The BamHI sites in pREMI-Z and at the integration sites in *H. polymorpha* sequences are indicated in white letters on a black background. Sequences in grey boxes indicate the expected position of the BamHI site of pREMI-Z in cases where no BamHI site was found in the sequences. Gaps were introduced to maximize the similarity

antibiotic G418 in some other organisms, zeocin may, when used in increasing amounts, select for multicopy integrations. However, no evidence supporting this view has been obtained yet. Multicopy integration is not a major disadvantage of the system, since simple molecular techniques allow the removal of excess copies of the plasmid prior to sequencing.

We observed that, in general, the pREMI-Z cassette integrates at a single locus in the *H. polymorpha* genome. Only in two transformants have we observed that the pREMI-Z cassette had possibly integrated at two different genomic loci. Although we have not analysed these mutants further, this type of integration requires the isolation of both genomic loci to determine which is linked to the observed mutant phenotype. Fortunately, the frequency of such events is low (<2.5%) and thus they will represent only a minor population of the mutants.

The ability of RALF to tag genes in *H. polymorpha* was clearly demonstrated by the isolation of auxotrophic mutants and mutants defective in peroxisome biogenesis (*pex*) and degradation (*pdd*). The *pex* mutants were isolated by making use of an eGFP-SKL reporter strain to select for Mut<sup>-</sup> mutants with altered peroxisomal profiles. The frequencies of Mut<sup>-</sup> and *pdd* mutants obtained with RALF were comparable to those obtained by the classical isolation procedures (Cregg et al. 1990; Titorenko et al. 1995). However, the RALF system permits direct identification of the mutagenised gene and should not generate the weak mutant phenotypes associated with many point mutations.

Sequence analysis of the rescued plasmids which contained pREMI-Z flanking genomic regions from *pex* mutants revealed in a number of cases *H. polymorpha* genes with a known role in peroxisome biogenesis or methanol metabolism. Although we identified some genes more than once, the integrations had occurred at different positions in the gene, arguing against the presence of recombinational hotspots in the *H. polymorpha* genome. The analysis of these mutants clearly demonstrates the feasibility of RALF in *H. polymorpha*. This is highlighted by the identification of a number of new genes whose function in peroxisome biogenesis is at present unknown. These mutants and their

corresponding genes are currently under investigation in our laboratory.

Sequences from a small percentage of the rescued plasmids showed identity of the flanking regions of pREMI-Z to two different DNA sequences. It is likely that in these mutants (a small) part of the genome has been deleted upon integration of the cassette. That deletions indeed occur was demonstrated by the isolation of the *PEX12* gene using the RVD3 mutant. So far, we have no data regarding the size of these deletions. The availability of the complete *H. polymorpha* genome sequence in the near future (Blandin et al. 2000) will provide conclusive evidence about the size of these deletions. However, we do not consider genomic deletions a serious problem because large numbers of mutants can be obtained rather easily in *H. polymorpha* and integration of the cassette was shown to occur at different places in the same gene. Therefore, mutants with genomic deletions can be omitted from further analysis. Alternatively, deleted regions of a gene can be isolated efficiently by PCR and/or functional complementation strategies.

In conclusion, the RALF procedure is a powerful tool that allows one successfully to tag, and rapidly isolate and identify, genes of interest in *H. polymorpha*. The integration is random and the high frequency of transformation of *H. polymorpha* using linear DNA fragments allows the screening of large numbers of mutants. Together with the complete genome sequence of *H. polymorpha*, which will be available shortly, this method provides us with a tool with which to identify rapidly genes of interest in our current research programme.

**Acknowledgements** We thank Arjen Krikken and Adriana Leao for information prior to publication. R.v.D. is supported by a grant from the Netherlands Technology Foundation (STW). K.N.F. is supported by an NWO-PULS Fellowship, J.A.K.W.K. by a grant from ALW, which is subsidised by the Dutch Organization for the Advancement of Pure Research (NWO).

## References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Baerends RJS, Faber KN, Kram AM, Kiel JAKW, van der Klei IJ, Veenhuis M (2000) A stretch of positively charged amino acids at the N terminus of *Hansenula polymorpha* Pex3p is involved in incorporation of the protein into the peroxisomal membrane. *J Biol Chem* 275:9986–9995
- Blandin G, Llorente B, Malpertuy A, Wincker P, Artiguenave F, Dujon B (2000) Genomic exploration of the hemiascomycetous yeasts. 13. *Pichia angusta*. *FEBS Lett* 487:76–81
- Bolker M, Bohnert HU, Braun KH, Gohl J, Kahmann R (1995) Tagging pathogenicity genes in *Ustilago maydis* by restriction enzyme-mediated integration (REMI). *Mol Gen Genet* 248:547–552
- Brown DH, Slobodkin IV, Kumamoto CA (1996) Stable transformation and regulated expression of an inducible reporter construct in *Candida albicans* using restriction enzyme-mediated integration. *Mol Gen Genet* 251:75–80

- Brown JS, Aufauvre-Brown A, Holden DW (1998) Insertional mutagenesis of *Aspergillus fumigatus*. *Mol Gen Genet* 259:327–335
- Cregg JM, van der Klei IJ, Sulter GJ, Veenhuis M, Harder W (1990) Peroxisome-deficient mutants of *Hansenula polymorpha*. *Yeast* 6:87–97
- Drocourt D, Calmels T, Reynes JP, Baron M, Tiraby G (1990) Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Res* 18:4009
- Faber KN, Swaving GJ, Faber F, AB G, Harder W, Veenhuis M, Haima P (1992) Chromosomal targeting of replicating plasmids in the yeast *Hansenula polymorpha*. *J Gen Microbiol* 138:2405–2416
- Faber KN, Haima P, Gietl C, Harder W, AB G, Veenhuis M (1994a) The methylotrophic yeast *Hansenula polymorpha* contains an inducible import pathway for peroxisomal matrix proteins with an N-terminal targeting signal (PTS2 proteins). *Proc Natl Acad Sci USA* 91:12985–12989
- Faber KN, Haima P, Harder W, Veenhuis M, AB G (1994b) Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. *Curr Genet* 25:305–310
- Gatignol A, Durand H, Tiraby G (1988) Bleomycin resistance conferred by a drug-binding protein. *FEBS Lett* 230:171–175
- Gleeson MAG, Sudbery PE (1988) Genetic analysis in the methylotrophic yeast *Hansenula polymorpha*. *Yeast* 4:293–303
- Granado JD, Kertesz-Chaloupkova K, Aebi M, Kues U (1997) Restriction enzyme-mediated DNA integration in *Coprinus cinereus*. *Mol Gen Genet* 256:28–36
- Kalish JE, Keller GA, Morrell JC, Mihalik SJ, Smith B, Cregg JM, Gould SJ (1996) Characterization of a novel component of the peroxisomal protein import apparatus using fluorescent peroxisomal proteins. *EMBO J* 15:3275–3285
- Kawabe Y, Enomoto T, Morio T, Urushihara H, Tanaka Y (1999) LbrA, a protein predicted to have a role in vesicle trafficking, is necessary for normal morphogenesis in *Polysphondylium pallidum*. *Gene* 239:75–79
- Kiel JAKW, Hilbrands RE, van der Klei IJ, Rasmussen SW, Salomons FA, van der Heide M, Faber KN, Cregg JM, Veenhuis M (1999) *Hansenula polymorpha* Pex1p and Pex6p are peroxisome-associated AAA proteins that functionally and physically interact. *Yeast* 15:1059–1078
- Kuspa A, Loomis WF (1992) Tagging developmental genes in *Dictyostelium* by restriction enzyme-mediated integration of plasmid DNA. *Proc Natl Acad Sci USA* 89:8803–8807
- Linnemannstons P, Voss T, Hedden P, Gaskin P, Tudzynski B (1999) Deletions in the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi* by restriction enzyme-mediated integration and conventional transformation-mediated mutagenesis. *Appl Environ Microbiol* 65:2558–2564
- Manoil C, Traxler B (2000) Insertion of in-frame sequence tags into proteins using transposons. *Methods* 20:55–61
- Salomons FA, Kiel JAKW, Faber KN, Veenhuis M, van der Klei IJ (2000) Overproduction of Pex5p stimulates import of alcohol oxidase and dihydroxyacetone synthase in a *Hansenula polymorpha* pex14 null mutant. *J Biol Chem* 275:12603–12611
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sanchez O, Navarro RE, Aguirre J (1998) Increased transformation frequency and tagging of developmental genes in *Aspergillus nidulans* by restriction enzyme-mediated integration (REMI). *Mol Gen Genet* 258:89–94
- Sato T, Yaegashi K, Ishii S, Hirano T, Kajiwarra S, Shishido K, Enei H (1998) Transformation of the edible basidiomycete *Lentinus edodes* by restriction enzyme-mediated integration of plasmid DNA. *Biosci Biotechnol Biochem* 62:2346–2350
- Schieftl RH, Petes TD (1991) Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 88:7585–7589
- Shuster JR, Bindel CM (1999) Promoter-tagged restriction enzyme-mediated insertion (PT-REMI) mutagenesis in *Aspergillus niger*. *Mol Gen Genet* 262:27–34
- Tan X, Waterham HR, Veenhuis M, Cregg JM (1995) The *Hansenula polymorpha* PER8 gene encodes a novel peroxisomal integral membrane protein involved in proliferation. *J Cell Biol* 128:307–319
- Titorenko VI, Keizer I, Harder W, Veenhuis M (1995) Isolation and characterization of mutants impaired in the selective degradation of peroxisomes in the yeast *Hansenula polymorpha*. *J Bacteriol* 177:357–363
- Van der Klei IJ, Harder W, Veenhuis M (1991) Methanol metabolism in a peroxisome-deficient mutant of *Hansenula polymorpha*: a physiological study. *Arch Microbiol* 156:15–23
- Van Dijk R, Faber KN, Kiel JAKW, Veenhuis M, van der Klei IJ (2000) The methylotrophic yeast *Hansenula polymorpha*: a versatile cell factory. *Enzyme Microb Technol* 26:793–800
- Van Dijken JP, Otto R, Harder W (1976) Growth of *Hansenula polymorpha* in a methanol-limited chemostat. Physiological responses due to the involvement of methanol oxidase as a key enzyme in methanol metabolism. *Arch Microbiol* 111:137–144
- Veenhuis M, Douma AC, Harder W, Osumi M (1983) Degradation and turnover of peroxisomes in the yeast *Hansenula polymorpha* induced by selective inactivation of peroxisomal enzymes. *Arch Microbiol* 134:193–203
- Veenhuis M, Salomons FA, van der Klei IJ (2000) Peroxisome biogenesis and degradation in yeast: a structure/function analysis. *Microsc Res Techn* 51:584–600
- Waterham HR, Titorenko VI, van der Klei IJ, Harder W, Veenhuis M (1992) Isolation and characterization of peroxisomal protein import (pim-) mutants of *Hansenula polymorpha*. *Yeast* 8:961–972